

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

This Page Blank (uspto)

①9 RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
PARIS

①1 N° de publication :
(à n'utiliser que pour les
commandes de reproduction)

2 679 083

②1 N° d'enregistrement national :

91 08607

⑤1 Int Cl⁵ : H 04 B 1/16

⑫

DEMANDE DE BREVET D'INVENTION

A1

②2 Date de dépôt : 09.07.91.

③0 Priorité :

④3 Date de la mise à disposition du public de la
demande : 15.01.93 Bulletin 93/02.

⑤6 Liste des documents cités dans le rapport de
recherche : *Se reporter à la fin du présent fascicule.*

⑥0 Références à d'autres documents nationaux
apparentés :

⑦1 Demandeur(s) : Société dite: THOMSON-CSF
(Société Anonyme) — FR.

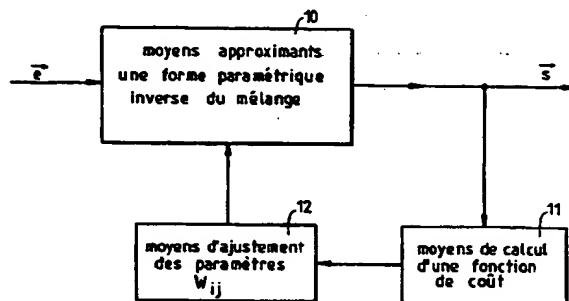
⑦2 Inventeur(s) : Burel Gilles.

⑦3 Titulaire(s) :

⑦4 Mandataire : Nguyen Dominique.

⑤4 Procédé et dispositif automatique de séparation de sources.

⑤7 Le procédé de séparation de sources comporte une phase préalable, pendant laquelle une mesure du degré de dépendance de n signaux est définie; une phase d'apprentissage pendant laquelle des exemples de mélange de n signaux e_i , à e_n sont utilisés pour ajuster, par approches successives, des paramètres W_i d'une forme paramétrique inverse associée à une transformation inverse, de façon à minimiser une fonction de coût déterminée à partir de la mesure du degré de dépendance de n signaux s_i , à s_n obtenus par transformation inverse d'un mélange donné; une phase de séparation de sources, pendant laquelle n signaux originaux nouveaux sont estimés à partir d'une observation d'un mélange instantané de ces signaux, en calculant la transformation inverse de l'observation.



FR 2 679 083 - A1



PROCÉDE ET DISPOSITIF AUTOMATIQUE DE SEPARATION DE SOURCES

L'invention se rapporte au domaine du traitement du signal et concerne un procédé automatique qui permet de retrouver n signaux originaux, appelés sources, à partir de l'observation d'un mélange instantané de ces signaux. Les applications concernent tous les domaines où l'on observe des mélanges de sources indépendantes, et où il est nécessaire de retrouver ces sources ; notamment, dans le domaine médical où les mesures fournies par les capteurs sont généralement des mélanges de signaux issus de plusieurs sources indépendantes et où le problème consiste à extraire de ces mélanges les signaux utiles. Le même problème se trouve également dans le domaine du traitement d'antenne pour la séparation de sources sonar ou radar, ou encore dans le domaine du traitement de la parole pour l'amélioration des signaux de diction où il est nécessaire d'éliminer des sources sonores parasites dues par exemple à la présence d'autres locuteurs ou de bruits de moteurs dans un avion. Des applications à la compression d'images peuvent être également envisagées, bien qu'il ne s'agisse pas d'un problème de séparation de sources.

Il existe différentes méthodes de séparation de sources, notamment des méthodes qui nécessitent des modèles a priori des sources, des méthodes qui cherchent à localiser les sources et des méthodes qui exploitent l'indépendance des sources.

Les méthodes qui exploitent l'indépendance des sources, se sont limitées au cas linéaire et consistent à rechercher une solution sous la forme suivante :

$$\begin{cases} \vec{e}(t) = A \vec{x}(t) \\ \vec{s}(t) = F \vec{e}(t) \end{cases}$$

où t , $\vec{x}(t)$, $\vec{e}(t)$, $\vec{s}(t)$ représentent respectivement le temps, les sources, les observations, les sorties d'un algorithme de recherche de la solution, dit algorithme d'apprentissage, et où A et F sont des matrices $n \times n$.

5 L'algorithme cherche à estimer la matrice F de façon que les sorties $\vec{s}(t)$ approchent au mieux les sources $\vec{x}(t)$; c'est-à-dire de façon que l'on retrouve dans $\vec{s}(t)$ les composantes de $\vec{x}(t)$, à une permutation et une dilatation près.

10 L'estimation de la matrice F peut être effectuée de différentes façons :

- une première méthode consiste à chercher une solution sous la forme $F = (I+G)^{-1}$ où I est la matrice identité, et G une matrice dont les termes diagonaux sont nuls. L'algorithme estime la matrice G par un processus itératif, puis la matrice F en est déduite. La matrice G qui réalise la séparation des sources est un état stable du processus itératif, mais le comportement de l'algorithme n'est pas connu lorsque la matrice G choisie comme solution initiale du processus itératif est éloignée de la véritable solution. L'apprentissage se décompose généralement en 2 phases : une première phase rapide qui aboutit à la décorrélation des sources, puis une phase plus lente qui réalise l'indépendance des sources,

25 - une deuxième méthode plus directe consiste à établir un système d'équations entre les cumulants d'ordre 4 de \vec{e} et ceux de \vec{s} (les cumulants étant des combinaisons de moments) dont les inconnues sont les coefficients de la matrice F . L'annulation des cumulants de \vec{s} est imposée puis le système d'équations est résolu à l'aide d'une méthode itérative de moindres carrés. Cette méthode semble plus rapide et robuste que la méthode précédente. En particulier, la rapidité est assurée par d'astucieuses décompositions matricielles.

30 Cependant, ces méthodes antérieures ne permettent de traiter que des mélanges linéaires, et de plus, certains aspects de ces méthodes ne sont pas clairement établis sur le plan théorique. Notamment, en ce qui concerne la première méthode

citée, le comportement de l'algorithme n'est pas connu lorsque la solution choisie pour initialiser le processus itératif est éloignée de la véritable solution au problème.

5 L'invention concerne un procédé de séparation de sources fondamentalement différent des procédés classiques qui n'est pas limité aux mélanges linéaires et qui ne nécessite a priori aucune connaissance sur les sources elles-mêmes, la seule contrainte étant que les sources soient indépendantes.

10 Selon l'invention, le procédé de séparation de sources permettant d'estimer n signaux originaux, appelés sources, à partir d'une observation d'un mélange instantané de ces signaux, les n sources étant indépendantes et de variance non nulle, l'observation étant constituée de n signaux, le mélange étant une transformation dont une forme paramétrique
15 est connue mais dont les paramètres sont inconnus, est caractérisé en ce qu'il comporte :

- une phase préalable, pendant laquelle une mesure du degré de dépendance de n signaux est définie,

- une phase d'apprentissage, pendant laquelle des
20 exemples de mélange de n signaux sont utilisés pour ajuster, par approches successives, des paramètres W_{ij} d'une forme paramétrique inverse associée à une transformation inverse, de façon à minimiser la mesure du degré de dépendance de n signaux obtenus par transformation inverse d'un mélange donné,

- une phase de séparation de sources, pendant
25 laquelle n signaux originaux nouveaux sont estimés à partir d'une observation d'un mélange instantané de ces signaux, en calculant la transformation inverse de l'observation.

30 D'autres particularités et avantages de l'invention apparaîtront clairement dans la description suivante donnée à titre d'exemple non limitatif et faite en regard des figures annexées qui représentent :

- la figure 1, un exemple de dispositif de séparation de sources, selon l'invention,

- la figure 2, un exemple de réseau de neurones multicouches,

- la figure 3, une représentation schématique de la structure interne d'un neurone,

5 - la figure 4, un organigramme de la phase d'apprentissage, selon l'invention.

Le problème de la séparation de sources consiste à estimer n signaux aléatoires ou déterministes x_1, x_2, \dots, x_n , à partir de l'observation d'une transformation instantanée, notée v , de ces signaux. Ces signaux sont appelés "sources" et sont notés par le vecteur $\vec{x}(t)$; les observations sont notées par le vecteur $\vec{e}(t)$; l'équation de la transformation instantanée s'écrit $\vec{e} = v(\vec{x})$; le paramètre t désigne le temps.

15 Pour résoudre ce problème, aucune connaissance a priori sur les sources elles-mêmes n'est utilisée, mais seules les trois hypothèses suivantes sont faites :

- les n sources sont indépendantes et de variance non nulle,

20 - l'observation est constituée de n signaux

e_1, e_2, \dots, e_n .

- la forme paramétrique de la transformation v est connue, mais les valeurs numériques des paramètres sont inconnues.

25 Ces hypothèses étant posées, la forme paramétrique inverse du mélange, notée w , est considérée. Cette forme paramétrique est en général connue car les processus physiques qui interviennent dans chaque application particulière envisagée sont connus (par exemple les phénomènes d'atténuation dans le domaine des radiocommunications). Par contre, les valeurs numériques des paramètres de cette forme paramétrique inverse sont inconnues. Il s'agit alors de déterminer les valeurs de ces paramètres, notés W_{ij} , de façon que les signaux, notés par le vecteur \vec{s} , obtenus en calculant la transformation inverse des observations, correspondent aux signaux sources \vec{x} à une

35

permutation et une déformation près. L'équation de la transformation inverse s'écrit $\vec{s} = w(\vec{e})$ et les composantes de \vec{s} doivent donc être égales aux composantes de \vec{x} à une permutation et une déformation près.

5 Pour que le vecteur \vec{s} soit une estimation du vecteur \vec{x} , il est nécessaire que les composantes de \vec{s} , notées s_i , soient indépendantes. Il s'agit donc de définir une mesure du degré de dépendance des composantes de \vec{s} et d'ajuster par approches successives les paramètres W_{ij} de la forme
10 paramétrique inverse w de façon à minimiser cette mesure du degré de dépendance des composantes de \vec{s} .

La mesure du degré de dépendance des composantes de \vec{s} sera appelée mesure de dépendance dans la suite de la description.

15 La première phase du procédé consiste à définir une mesure de dépendance des composantes s_i de \vec{s} . Les phénomènes physiques intervenant dans le problème de la séparation de sources étant des phénomènes stochastiques donc soumis aux lois du hasard, une mesure de dépendance peut être
20 définie à partir des densités de probabilité $p_{S_i}(s_i)$ des composantes s_i de \vec{s} , où S_i est une variable aléatoire et s_i est une valeur particulière de cette variable aléatoire. Par conséquent, une condition nécessaire et suffisante pour que les composantes s_i soient indépendantes est donnée par
25 l'équation (1) suivante :

$$(1) \quad p_{s_1 \dots s_n}(s_1, \dots, s_n) = p_{S_1}(s_1) \dots p_{S_n}(s_n)$$

Lorsque les composantes s_i sont indépendantes, la mesure de dépendance doit être nulle. Par conséquent,
30 l'expression (2) suivante peut servir de base pour choisir une mesure de dépendance :

$$(2) \quad p_{s_1 \dots s_n}(s_1, \dots, s_n) - p_{S_1}(s_1) \dots p_{S_n}(s_n)$$

En fait, les composantes s_i doivent être
35 indépendantes en moyenne ; par conséquent l'expression (2) est

intégrée sur toutes les valeurs possibles de s_i . En outre, pour obtenir une mesure de dépendance bien adaptée au traitement de signaux réels, c'est-à-dire pour éviter les variations brutales de la mesure de dépendance, pour atténuer des erreurs d'arrondi ou de mesure, et pour éviter également la divergence de l'intégrale lorsque les composantes s_i ont des valeurs discrètes, un filtrage gaussien est introduit sur le terme à intégrer. La mesure de dépendance est alors donnée par l'expression (3) suivante :

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \left([p_{S_1 \dots S_n}(s_1, \dots, s_n) - p_{S_1}(s_1) \dots p_{S_n}(s_n)] * \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{s_i^2}{2\sigma_i^2}} \right)^2 ds_1 \dots ds_n$$

où les σ_i sont les écarts types des composantes s_i et où le signe "*" représente un produit de convolution.

Etant donné qu'il est difficile d'exploiter une telle expression car il serait nécessaire d'estimer les densités de probabilité des composantes s_i et de réaliser une intégration, cette expression est mise sous une forme équivalente en effectuant tout d'abord une transformation de Fourier.

La transformée de Fourier de la densité de probabilité P_{S_i} est aussi la fonction caractéristique, notée Φ_{S_i} , telle que

$$(4) \quad \Phi_{S_1 \dots S_n}(u_1, \dots, u_n) = E \left\{ e^{-j(S_1 u_1 + \dots + S_n u_n)} \right\}$$

où E désigne la moyenne arithmétique, et u est une variable certaine.

Compte tenu du théorème de Plancherel qui exprime le fait que la transformation de Fourier conserve la norme, et compte tenu des propriétés de la transformation de Fourier, l'expression de la mesure de dépendance s'écrit aussi sous la forme (5) suivante :

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} | \Phi_{S_1 \dots S_n}(u_1, \dots, u_n) - \Phi_{S_1}(u_1) \dots \Phi_{S_n}(u_n) |^2 e^{-\sigma_1^2 u_1^2} \dots e^{-\sigma_n^2 u_n^2} du_1 \dots du_n$$

où $\Phi_{S_1 \dots S_n}(u_1, \dots, u_n)$ est la transformée de Fourier de $P_{S_1 \dots S_n}(s_1, \dots, s_n)$.

5 Etant donné que les phénomènes physiques qui interviennent dans les problèmes de séparation de sources sont des phénomènes réels, donc bornés, et que les moments de valeur absolue bornée existent, il est possible d'expliciter la mesure de dépendance en décomposant les fonctions caractéristiques Φ_{s_i} en série de Mac Laurin de façon à faire apparaître les moments des composantes s_i .

10 La mesure de dépendance est alors donnée par l'équation (6) suivante :

$$(6) \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \left| \sum_{\alpha_1 \dots \alpha_n} T_{\alpha_1 \dots \alpha_n} u_1^{\alpha_1} \dots u_n^{\alpha_n} \right|^2 e^{-\sigma_1^2 u_1^2} \dots e^{-\sigma_n^2 u_n^2} du_1 \dots du_n$$

15

$$\text{avec } T_{\alpha_1 \dots \alpha_n} = \frac{1}{\alpha_1! \dots \alpha_n!} \left\{ \frac{\partial^{\alpha_1 + \dots + \alpha_n} \Phi_{S_1 \dots S_n}}{\partial u_1^{\alpha_1} \dots \partial u_n^{\alpha_n}}(0, \dots, 0) - \frac{\partial^{\alpha_1} \Phi_{S_1}}{\partial u_1^{\alpha_1}}(0) \dots \frac{\partial^{\alpha_n} \Phi_{S_n}}{\partial u_n^{\alpha_n}}(0) \right\}.$$

où α_i sont des coefficients qui indiquent l'ordre des dérivées partielles. En pratique, le développement en série de Mac Laurin peut être limité à l'ordre K, par exemple égal à 5, les termes suivants étant très faibles.

20 En se référant à l'équation (4) et en notant M_{α_i} le moment d'ordre α_i de la composante s_i , $M_{\alpha_1 \dots \alpha_n}$ étant les moments croisés d'ordre $\alpha_1, \dots, \alpha_n$ des composantes s_1, \dots, s_n , $T_{\alpha_1 \dots \alpha_n}$ s'écrit aussi :

25

$$(7) \quad T_{\alpha_1 \dots \alpha_n} = \frac{1}{\alpha_1! \dots \alpha_n!} (-j)^{\alpha_1 + \dots + \alpha_n} M_{\alpha_1 \dots \alpha_n}$$

30 en développant dans l'équation (6) la norme et en utilisant la linéarité de l'intégration, une expression de la mesure de dépendance est obtenu conformément à l'équation (8) suivante :

$$(8) \quad \sum_{\substack{\alpha_1 \dots \alpha_n \\ \alpha_1 + \dots + \alpha_n < K}} \sum_{\substack{\beta_1 \dots \beta_n \\ \beta_1 + \dots + \beta_n < K}} H_{\alpha_1 \dots \alpha_n \beta_1 \dots \beta_n} M_{\alpha_1 \dots \alpha_n} M_{\beta_1 \dots \beta_n}$$

35

Les coefficients α_i et β_i sont des nombres entiers positifs ou nuls. K est un nombre entier positif qui correspond au dernier ordre de développement de l'équation (5) en série de Mac Laurin. Les coefficients $H_{\alpha_1 \dots \alpha_n, \beta_1 \dots \beta_n}$ sont des nombres réels tels que :

$$H_{\alpha_1 \dots \alpha_n, \beta_1 \dots \beta_n} = \prod_{i=1}^n g(\alpha_i, \beta_i, \sigma_i)$$

10

où
$$g(\alpha, \beta, \sigma) = \frac{1}{\alpha! \beta!} (-1)^{\frac{\alpha-\beta}{2}} J_{\alpha+\beta}(\sigma) \quad \begin{array}{l} \text{si } \alpha + \beta \text{ est pair} \\ = 0 \quad \quad \quad \text{sinon} \end{array}$$

15

$$J_{2k}(\sigma) = \frac{(2k)!}{4^k k!} \frac{\sqrt{2\pi}}{\sigma^{2k-1}}$$

20

Lorsque l'un au moins des indices α_i ou β_i tend vers l'infini, le coefficient $H_{\alpha_1 \dots \alpha_n, \beta_1 \dots \beta_n}$ tend vers zéro. Ceci permet de limiter la sommation de l'expression (8) à un ordre K .

25

L'expression (8) de la mesure de dépendance en fonction des moments est une forme quadratique définie positive. Elle n'est donc nulle que lorsque tous les moments $M_{\alpha_1 \dots \alpha_n}$ sont nuls. Pour des valeurs de K tendant vers l'infini, cette expression (8) est strictement équivalente à l'expression (3).

30

Etant donné le nombre important de calculs à réaliser, en pratique il est souvent possible de remplacer cette expression (8) par une forme quadratique définie positive plus simple. Notamment, pour un mélange comportant un nombre L

35

d'inconnues, une forme quadratique définie positive conduisant à annuler L moments peut être suffisante, car L équations sont utilisées. Mais la non redondance de ces équations n'est nullement garantie, et en cas d'échec de la séparation, il faut se rapprocher de la forme définie par l'équation (8) dont l'annulation garantit toujours l'indépendance, quels que soient le mélange et les signaux sources.

La deuxième phase du procédé est une phase d'apprentissage qui consiste à ajuster, par approches successives, des paramètres W_{ij} d'une structure approximant la forme paramétrique inverse du mélange de signaux de façon à minimiser une fonction, appelée fonction de coût, qui représente une mesure de dépendance des composantes s_i de \vec{s} .

La figure 1 représente un exemple de dispositif de séparation de sources.

Ce dispositif comprend des moyens 10 approximant une forme paramétrique inverse d'un mélange de n signaux, des moyens 11 de calcul d'une fonction de coût et des moyens 12 d'ajustement des paramètres W_{ij} de la forme paramétrique inverse. Les moyens 10 reçoivent en entrée n signaux e_1 à e_n provenant de l'observation d'une transformation instantanée de n signaux sources, et délivrent en sortie n signaux s_1 à s_n correspondant à une transformation inverse des signaux e_1 à e_n .

Les moyens 11 reçoivent en entrée les n signaux s_1 à s_n provenant des moyens 10 et calculent la fonction de coût correspondant à ces signaux, cette fonction de coût étant une mesure de dépendance des signaux s_1 à s_n .

Les moyens 12 reçoivent en entrée la valeur de la fonction de coût calculée par les moyens 11 et calculent de nouvelles valeurs des paramètres W_{ij} de façon à minimiser la fonction de coût. Ces nouvelles valeurs des paramètres W_{ij} sont transmises en retour aux moyens 10.

Les réseaux de neurones multicouches sont des structures suffisamment générales pour représenter, ou au moins

approximer, un grand nombre de formes paramétriques, les poids synaptiques du réseau étant les paramètres W_{ij} . Par conséquent, dans l'exemple décrit précisément ci-après, c'est un réseau de neurones multicouches qui est utilisé pour approximer la forme paramétrique inverse du mélange de signaux. Le dimensionnement du réseau ainsi que le choix des fonctions d'activation des neurones sont effectués de façon à être conformes à la forme paramétrique inverse du mélange de signaux. Le vecteur d'entrée du réseau de neurones est noté \vec{e} et correspond à l'observation d'un mélange instantané de n signaux originaux constituant les sources notées par le vecteur \vec{x} , le vecteur de sortie du réseau de neurones est noté \vec{s} et correspond à une estimation du vecteur \vec{x} à une permutation et à une déformation près. Les vecteurs \vec{e} , \vec{s} et \vec{x} ont le même nombre de composantes et ont donc la même dimension.

La figure 2 représente un exemple de réseau de neurones multicouches.

Ce réseau de neurones est organisé en plusieurs niveaux, les sorties d'un niveau constituant les entrées du niveau suivant. Ce réseau comporte trois niveaux comportant respectivement N_1 , N_2 , N_3 neurones. Les neurones du premier niveau n'effectuent aucun traitement et transmettent simplement la valeur e_i qui leur est présentée.

Les premier et deuxième niveaux comportent en outre un neurone seuil NS qui n'est pas connecté aux entrées et qui permet d'ajouter un degré de liberté supplémentaire dans la masse d'informations que constituent les poids synaptiques du réseau de neurones. Ce degré de liberté supplémentaire permet au réseau d'apprendre plus facilement et plus vite.

Dans le cas du problème de la séparation de sources, les première et dernière couches du réseau de neurones comportent le même nombre de neurones. Ce nombre de neurones est égal à la dimension du vecteur \vec{e} et à celle du vecteur \vec{s} .

La figure 3 représente un exemple de modélisation d'un neurone. Un neurone J peut être modélisé à l'aide d'un

sommateur pondéré, 30, associé à un dispositif de calcul d'une fonction non linéaire, 31. Les valeurs aux entrées $I_1, I_2, \dots, I_i, \dots, I_n$ du sommateur pondéré, 30, sont multipliées par des coefficients de pondération respectivement $W_{1j}, W_{2j}, \dots, W_{ij}, \dots, W_{nj}$ puis additionnées pour former une valeur intermédiaire a_j appelée potentiel du neurone J à laquelle est appliqué un calcul de fonction correspondant, $f(a_j)$, qui fournit la valeur de la sortie O_j de ce neurone J. La fonction $f(a_j)$ est aussi appelée fonction d'activation des neurones. Les coefficients de pondération sont les poids synaptiques W_{ij} des connexions neuronales entre un neurone J et les entrées e_i .

La phase d'apprentissage est effectuée de la manière suivante décrite ci-après en référence à la figure 4.

Dans une première étape 1 les poids synaptiques W_{ij} des connexions neuronales entre un neurone J et les entrées e_i sont initialisés pour chacun des neurones. Les valeurs initiales des poids des connexions neuronales sont fixées aléatoirement selon une loi de probabilité uniforme dans un intervalle $-D, +D$. Pour les neurones d'un niveau donné, la valeur de D est choisie égale à $1/\sqrt{N}$, où N est le nombre de neurones du niveau inférieur.

Dans une étape 2, à un rang d'itération p donné, un exemple de mélange \vec{e} de n signaux est présenté aux entrées du réseau de neurones et les sorties s_i du réseau de neurones sont calculées. Le nombre d'itérations total peut être fixé à l'avance, par exemple p varie entre 1 et 10 000.

Dans une étape 3, une fonction de coût qui représente une mesure de dépendance des composantes s_i de \vec{s} est calculée. La fonction de coût choisie est la suivante :

$$C = \frac{1}{2} \left\{ \sum_{i=1}^n |E\{S_i\}|^2 + \sum_{i=1}^n |E\{S_i^2\} - 1|^2 + \sum_{\alpha_1 \dots \alpha_n, \beta_1 \dots \beta_n} H_{\alpha_1 \dots \alpha_n, \beta_1 \dots \beta_n} M_{\alpha_1 \dots \alpha_n} M_{\beta_1 \dots \beta_n} \right\}$$

où $E\{S_i\}$ et $E\{S_i^2\}$ représentent, respectivement, les valeurs moyennes de S_i et de S_i^2 .

Les coefficients $H_{\alpha_1 \dots \alpha_n, \beta_1 \dots \beta_n}$ peuvent être les coefficients déterminés ci-dessus ou peuvent avoir d'autres valeurs si l'on souhaite une expression simplifiée, mais il faut nécessairement que la fonction de coût soit une forme quadratique définie positive et que le nombre de moments affectés d'un coefficient non nul soit au moins égal au nombre d'inconnues dans le mélange.

Pour chaque composante s_i de \vec{s} , cette fonction de coût, C , est la somme de trois termes différents. Le choix de cette fonction de coût permet d'imposer trois contraintes différentes pour que la séparation des sources soit réalisée.

Le premier terme de la fonction de coût impose que la valeur moyenne de s_i soit nulle pour toutes les sorties du réseau de neurones. Cette contrainte permet de simplifier les calculs et notamment de simplifier l'expression de la mesure de dépendance. Cette contrainte est toujours possible à réaliser lorsque la couche de neurones inférieure à la couche de sortie comporte un neurone seuil NS ; il suffit alors d'ajuster les poids synaptiques des connexions neuronales entre chaque neurone de la couche de sortie et le neurone seuil de la couche de neurones inférieure à cette couche de sortie.

Le deuxième terme de la fonction de coût impose que la variance de s_i soit égale à 1 pour toutes les sorties du réseau de neurones. Cette contrainte est indispensable car elle permet d'écarter la solution la plus simple pour réaliser l'indépendance des sources, c'est-à-dire celle qui consiste à obtenir des sorties identiquement nulles.

Le troisième terme de la fonction de coût impose que l'expression de la mesure de dépendance, (8), précédemment déterminée, soit nulle.

Les deux premières contraintes mentionnées ci-dessus permettent également de faciliter le choix des écarts types σ_i du filtrage gaussien qui intervient dans les coefficients $H_{\alpha_1 \dots \alpha_n, \beta_1 \dots \beta_n}$ de l'expression (8) de la mesure de dépendance. En effet, les valeurs de sorties recherchées étant centrées et

de variance unité, il suffit que tous les écarts types σ_i soient identiques car toutes les sorties tendent à avoir la même dispersion. Pour faciliter l'apprentissage, la puissance du filtrage gaussien peut être réduite progressivement. A titre d'exemple, l'apprentissage peut commencer avec une valeur de σ_i égale à 0,5 qui peut être réduite progressivement jusqu'à une valeur de σ_i égale à 0,05.

Dans une étape 4, les valeurs des paramètres W_{ij} sont modifiées en tenant compte des modifications des paramètres ΔW_{ij} (p-1), effectuées au rang d'itération précédent, p-1, ces valeurs étant pondérées par un coefficient de filtrage γ , et en tenant compte de la dérivée de la fonction de coût C par rapport à W_{ij} , cette dérivée étant pondérée par une valeur d'un paramètre η , dit "vitesse d'apprentissage", de la manière suivante :

$$(10) \quad \Delta W_{ij} (p) = - \eta \frac{\partial C}{\partial W_{ij}} (p) + \gamma \Delta W_{ij} (p-1)$$

Lorsque les paramètres W_{ij} sont modifiés, un test est effectué, dans une étape 5, pour voir s'il s'agit du dernier exemple. Si ce test est positif, l'apprentissage est terminé, si ce test est négatif, un nouvel exemple \vec{e} de n signaux est présenté aux entrées du réseau de neurones et les opérations des étapes 2 à 5 sont de nouveau effectuées de façon à minimiser la fonction de coût. Les paramètres d'apprentissage η et de filtrage γ sont choisis par l'opérateur, par exemple il pourra choisir $\eta = 0,005$ et $\gamma = 0,985$. Il est également possible de choisir une valeur de vitesse d'apprentissage plus élevée pour les premiers exemples et de diminuer progressivement cette valeur en cours d'apprentissage.

Pour calculer les modifications des paramètres ΔW_{ij} , il faut déterminer le gradient de la fonction de coût C par rapport aux poids synaptiques du réseau.

Pour déterminer ce gradient, il faut distinguer entre deux types différents d'apprentissage : l'apprentissage continu et l'apprentissage global.

L'apprentissage continu est nécessaire lorsque les exemples de mélange de signaux sont des échantillons $\vec{e}(t)$ présentés en temps réel aux entrées du réseau de neurones. Dans ce cas, les valeurs moyennes, à l'instant t , des composantes s_i de $\vec{s}(t)$ ne peuvent pas être calculées de façon exacte mais seulement estimées au moyen de filtres passe-bas de coefficient noté b , b étant un nombre réel dont la valeur est légèrement inférieure à 1. A titre d'exemple, la valeur de b peut être choisie égale à 0,98.

en notant $R_{\alpha_1 \dots \alpha_n} = E \{ S_1^{\alpha_1} \dots S_n^{\alpha_n} \}$

et $\vec{\Theta}_i = \underbrace{[0, \dots, 0, 1, 0, \dots, 0]}_n^{\perp}$

où E représente la valeur moyenne, les coefficients α_1 à α_n sont des nombres entiers positifs ou nuls, et le signe \perp représente la transposée matricielle, l'équation des filtres et l'équation des moments s'écrivent :

$$(11) \quad \hat{R}_{\alpha_1 \dots \alpha_n}(t) = (1 - b) S_1^{\alpha_1}(t) \dots S_n^{\alpha_n}(t) + b \hat{R}_{\alpha_1 \dots \alpha_n}(t-1)$$

$$(12) \quad \hat{M}_{\alpha_1 \dots \alpha_n}(t) = \hat{R}_{\alpha_1 \dots \alpha_n}(t) - \hat{R}_{\alpha_1 \Theta_1}(t) \dots \hat{R}_{\alpha_n \Theta_n}(t)$$

$\hat{R}_{\alpha_1 \dots \alpha_n}$ et $\hat{M}_{\alpha_1 \dots \alpha_n}$ étant respectivement des estimations de $R_{\alpha_1 \dots \alpha_n}$ et de $M_{\alpha_1 \dots \alpha_n}$ obtenues après filtrage.

Le gradient de la fonction coût peut être écrit de la façon suivante :

$$(13) \quad \frac{\partial \hat{C}(t)}{\partial W_{ik}} = \sum_{j=1}^n \frac{\partial \hat{C}(t)}{\partial S_j(t)} \frac{\partial S_j(t)}{\partial W_{ik}}$$

Le terme $\frac{\partial S_j}{\partial W_{ik}}(t)$ représente la

dérivée des sorties du réseau de neurones par rapport aux poids synaptiques et peut être calculé en utilisant un algorithme connu sous le nom d'algorithme de rétropropagation du gradient.

Le calcul du terme $\partial C(t) / \partial S_j(t)$ doit être effectué séparément ; son expression se met sous la forme (14) suivante :

$$\begin{aligned} \frac{\partial \hat{C}(t)}{\partial S_j(t)} = & (1 - b) \left\{ \hat{R}_{\theta_j}(t) + 2 \left(\hat{R}_{2\theta_j}(t) - 1 \right) S_j(t) \right\} \\ & + (1 - b) \left\{ \sum_{\substack{\alpha_1 \dots \alpha_n \\ \alpha_j > 0}} \alpha_j S_j^{\alpha_j - 1}(t) \left(\prod_{\substack{i=1 \\ i \neq j}}^n S_i^{\alpha_i}(t) - \prod_{\substack{i=1 \\ i \neq j}}^n \hat{R}_{\alpha_i \theta_i}(t) \right) \right. \\ & \left. \left(\sum_{\beta_1 \dots \beta_n} H_{\alpha_1 \dots \alpha_n \beta_1 \dots \beta_n} \hat{M}_{\beta_1 \dots \beta_n}(t) \right) \right\} \end{aligned}$$

15

L'apprentissage global peut être effectué lorsque les exemples présentés aux entrées du réseau de neurones sont disponibles et représentent un ensemble fini d'échantillons. Les moyennes des composantes s_j de \vec{s} peuvent alors être calculées par une sommation sur tous les échantillons possibles dont on dispose. En notant T le nombre d'échantillons, l'équation (15) de la valeur moyenne est la suivante :

$$(15) \quad \hat{R}_{\alpha_1 \dots \alpha_n} = \frac{1}{T} \sum_{t=1}^T S_1^{\alpha_1}(t) \dots S_n^{\alpha_n}(t)$$

25

Le gradient de la fonction de coût peut alors être écrit de la façon suivante :

30

$$(16) \quad \frac{\partial \hat{C}}{\partial W_{ik}} = \sum_{t=1}^T \sum_{j=1}^n \frac{\partial \hat{C}}{\partial S_j(t)} \frac{\partial S_j(t)}{\partial W_{ik}}$$

35

Le terme $\partial S_j(t) / \partial W_{ik}$ peut être calculé de la même façon que pour l'apprentissage continu, c'est-à-dire en utilisant l'algorithme de rétropropagation du gradient. Le

calcul du terme $\partial C / \partial S_j(t)$ doit être effectué séparément ;
son expression se met sous la forme (17) suivante :

$$\begin{aligned} \frac{\partial \hat{C}}{\partial S_j(t)} = & \frac{1}{T} \{ \hat{R}_{\theta_j} + 2 (\hat{R}_{2\theta_j} - 1) S_j(t) \} \\ & + \frac{1}{T} \left\{ \sum_{\substack{\alpha_1 \dots \alpha_n \\ \alpha_j > 0}} \alpha_j S_j^{\alpha_j - 1}(t) \left(\prod_{\substack{l=1 \\ l \neq j}}^n S_l^{\alpha_l}(t) - \prod_{\substack{l=1 \\ l \neq j}}^n \hat{R}_{\alpha_l \theta_l} \right) \left(\sum_{\beta_1 \dots \beta_n} H_{\alpha_1 \dots \alpha_n, \beta_1 \dots \beta_n} \hat{M}_{\beta_1 \dots \beta_n} \right) \right\} \end{aligned}$$

10

Dans le cas de l'apprentissage global, les exemples sont présentés aux entrées du réseau de neurones dans un ordre aléatoire, et un même exemple peut être présenté plusieurs fois.

15

A la fin de la phase d'apprentissage, lorsque tous les exemples disponibles ont été présentés au réseau de neurones, ou lorsque le nombre d'itération total, préalablement fixé, a été atteint, la fonction de coût est minimisée et les paramètres W_{ij} du réseau de neurones sont figés. Le réseau de neurones est alors prêt à réaliser la séparation de nouveaux mélanges de n sources indépendantes lorsque la forme paramétrique inverse de ces mélanges est la même que celle déterminée pendant la phase d'apprentissage.

20

Dans une troisième phase appelée phase de séparation de sources, les n signaux originaux correspondant à ces nouveaux mélanges sont estimés en calculant la transformation inverse de ces mélanges, cette transformation inverse étant effectuée par le réseau de neurones tel que configuré pendant la phase d'apprentissage.

25

30

La présente invention n'est pas limitée à l'exemple précisément décrit, notamment, la forme paramétrique inverse du mélange peut être approximée par un dispositif ayant une structure différente d'un réseau de neurones.

35

REVENDEICATIONS

1 - Procédé automatique de séparation de sources permettant d'estimer n signaux originaux, appelés sources, à partir d'une observation d'un mélange instantané de ces signaux, les n sources étant indépendantes et de variance non nulle, l'observation étant constituée de n signaux, le mélange étant une transformation dont une forme paramétrique est connue mais dont les paramètres sont inconnus, caractérisé en ce qu'il comporte :

- une phase préalable, pendant laquelle une mesure du degré de dépendance de n signaux est définie,

- une phase d'apprentissage, pendant laquelle des exemples de mélange de n signaux e_1 à e_n sont utilisés pour ajuster, par approches successives, des paramètres W_{ij} d'une forme paramétrique inverse associée à une transformation inverse, de façon à minimiser une fonction de coût déterminée à partir de la mesure du degré de dépendance de n signaux s_1 à s_n obtenus par transformation inverse d'un mélange donné,

- une phase de séparation de sources, pendant laquelle n signaux originaux nouveaux sont estimés à partir d'une observation d'un mélange instantané de ces signaux, en calculant la transformation inverse de l'observation.

2 - Procédé selon la revendication 1, caractérisé en ce que, les signaux sources étant des signaux aléatoires ou déterministes, la mesure du degré de dépendance des n signaux s_1 à s_n est une forme quadratique définie positive obtenue à partir des densités de probabilité $p_{S_i}(s_i)$, pour i variant de 1 à n , et dont une expression en fonction des moments croisés d'ordre $\alpha_1 \dots \alpha_n$ et d'ordre $\beta_1 \dots \beta_n$ des signaux s_1, \dots, s_n est la suivante :

$$\sum_{\substack{\alpha_1 \dots \alpha_n \\ \alpha_1 + \dots + \alpha_n < K}} \sum_{\substack{\beta_1 \dots \beta_n \\ \beta_1 + \dots + \beta_n < K}} H_{\alpha_1 \dots \alpha_n \beta_1 \dots \beta_n} M_{\alpha_1 \dots \alpha_n} M_{\beta_1 \dots \beta_n}$$

où les coefficients $H_{\alpha_1 \dots \alpha_n, \beta_1 \dots \beta_n}$ sont des nombres réels, et K est un nombre entier positif.

3 - Procédé selon la revendication 2, caractérisé en ce que la fonction de coût est choisie de façon à imposer trois contraintes différentes pour que la séparation de sources soit réalisée, les deux premières contraintes imposent que les signaux s_i aient une valeur moyenne nulle et une variance égale à l'unité, et la troisième contrainte impose que la mesure du degré de dépendance soit nulle.

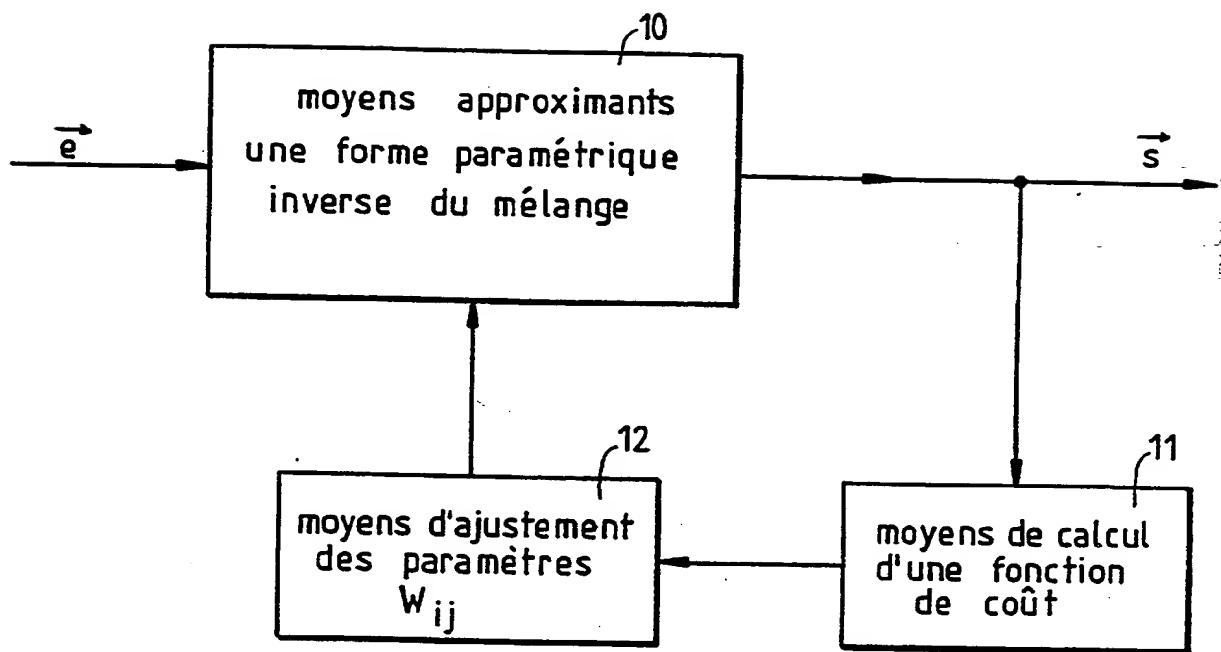
4 - Dispositif de séparation de sources pour la mise en oeuvre du procédé selon l'une quelconque des revendications précédentes, caractérisé en ce qu'il comporte :

- des moyens (10) approximant une forme paramétrique inverse du mélange destinés à recevoir, en entrée, n signaux e_1 à e_n provenant de l'observation d'une transformation instantanée de n signaux sources, et à délivrer en sortie n signaux s_1 à s_n correspondant à une transformation inverse des signaux e_1 à e_n ,

- des moyens (11) connectés en sortie des moyens (10) et destinés à calculer la fonction de coût C correspondant aux signaux s_1 à s_n ,

- des moyens (12) connectés en sortie des moyens (10) et destinés à calculer des valeurs des paramètres W_{ij} de la forme paramétrique inverse du mélange qui minimisent la fonction de coût C, et à transmettre en retour, aux moyens (10), ces valeurs des paramètres W_{ij} .

FIG. 1



2/3

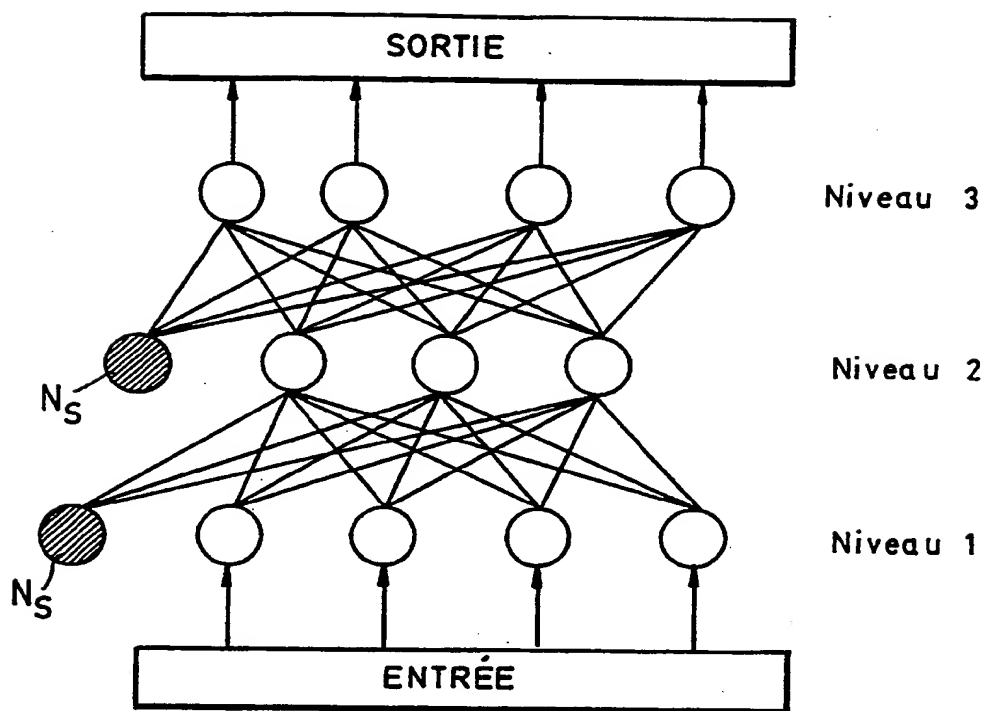


FIG. 2

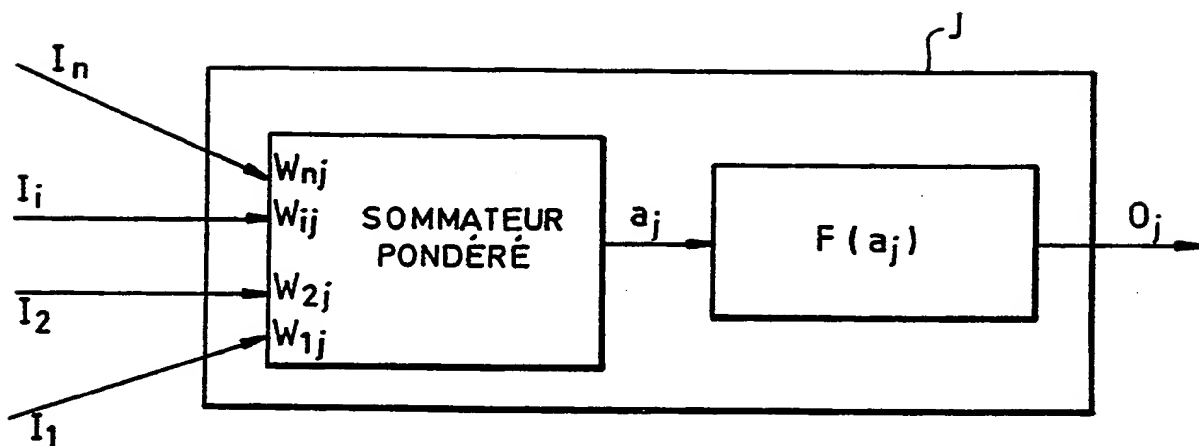


FIG. 3

3/3

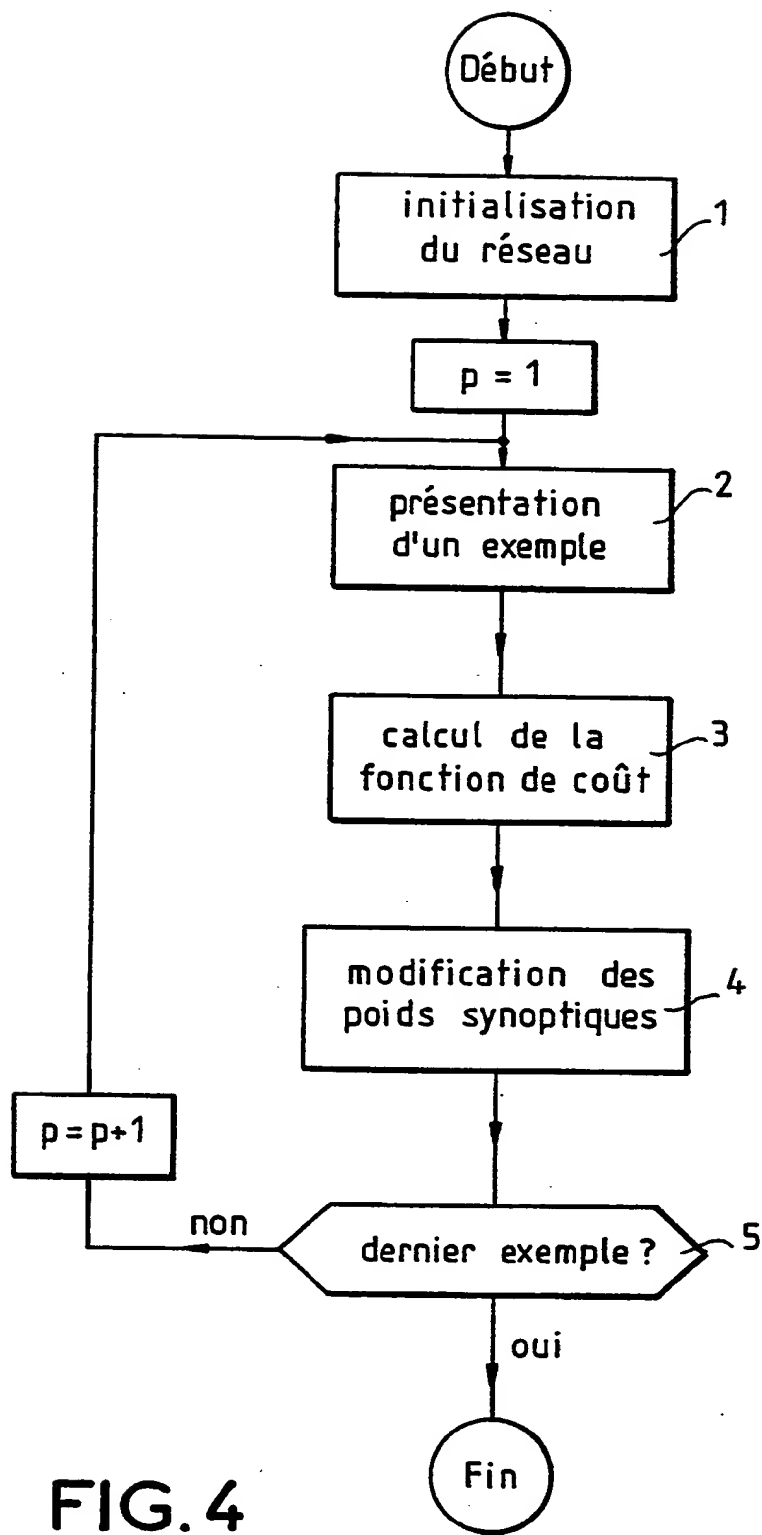


FIG. 4

INSTITUT NATIONAL
de la
PROPRIETE INDUSTRIELLERAPPORT DE RECHERCHE
établi sur la base des dernières revendications
déposées avant le commencement de la rechercheFR 9108607
FA 464862

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
X	TRAITEMENT DU SIGNAL vol. 5, no. 6, 1988, pages 389 - 403; JUTTEN: 'Une solution neuromimetique au probleme de separation de sources' * page 390, colonne de gauche, ligne 1 - page 402, colonne de gauche, ligne 59; figures 1-10 *	1-4
X	ICASSP 89 : INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING vol. S1, 23 Mai 1989, GLASGOW, SCOTLAND pages 357 - 360; ROGERS: 'Neural network enhancement for a two speaker separation system' * page 357, colonne de gauche, ligne 1 - page 360, colonne de gauche, ligne 8; figures 1-6 *	1,4
A	WO-A-9 102 324 (HUGHES AIRCRAFT COMPANY) 21 Février 1991 * page 1, ligne 9 - page 3, ligne 17 * * page 4, ligne 32 - page 5, ligne 12 * * page 6, ligne 15 - page 13, ligne 35; figures 1,4 *	2,3
X	WO-A-9 102 324 (HUGHES AIRCRAFT COMPANY) 21 Février 1991 * page 1, ligne 9 - page 3, ligne 17 * * page 4, ligne 32 - page 5, ligne 12 * * page 6, ligne 15 - page 13, ligne 35; figures 1,4 *	1,4
A	SIGNAL PROCESSING. vol. 24, no. 1, Juillet 1991, AMSTERDAM NL pages 1 - 10; JUTTEN: 'Blind separation of sources, Part 1 : an adaptive algorithm based on neuromimetic architecture' * page 1, colonne de gauche, ligne 1 - page 2, colonne de droite, ligne 15 *	2,3
A	SIGNAL PROCESSING. vol. 24, no. 1, Juillet 1991, AMSTERDAM NL pages 1 - 10; JUTTEN: 'Blind separation of sources, Part 1 : an adaptive algorithm based on neuromimetic architecture' * page 1, colonne de gauche, ligne 1 - page 2, colonne de droite, ligne 15 *	1
Date d'achèvement de la recherche 03 AVRIL 1992		Examinateur SCHENKELS P. F.
<p>CATEGORIE DES DOCUMENTS CITES</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général O : divulgation non-écrite P : document intercalaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant</p>		

This Page Blank (uspto)